



TSM-33

JPW

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: M. ICHIKAWA, et al
Serial No.: 10/608,010
Filing Date: June 30, 2003
For: HOT STANDBY SERVER SYSTEM
Art Unit: 2113
Examiner: A. Riad

LETTER CLAIMING RIGHT OF PRIORITY

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

June 14, 2006

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, applicants hereby claim
the right of priority based on:

**Japanese Application No. 2002-266206
Filed: September 12, 2002**

A Certified copy of said application document is attached hereto.

Respectfully submitted,

Carl I. Brundidge
Registration No. 29,621
MATTINGLY, STANGER, MALUR & BRUNDIDGE, P.C.

CIB/jdc
Enclosures
703/684-1120

日 本 国 特 許 庁

JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2002年 9月12日

出 願 番 号

Application Number:

特願2002-266206

[ST.10/C]:

[JP2002-266206]

出 願 人

Applicant(s):

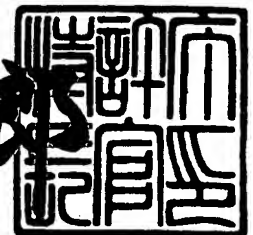
株式会社日立製作所

CERTIFIED COPY OF
PRIORITY DOCUMENT

2003年 5月30日

特 許 庁 長 官
Commissioner,
Japan Patent Office

太田信一



【書類名】 特許願

【整理番号】 K02009851A

【あて先】 特許庁長官殿

【国際特許分類】 G06F 11/20

【発明者】

 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内

 【氏名】 市川 正也

【発明者】

 【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所ソフトウェア事業部内

 【氏名】 三瓶 英智

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

 【氏名】 鶴飼 敏之

【発明者】

 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内

 【氏名】 春名 高明

【発明者】

 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内

 【氏名】 二瀬 健太

【発明者】

 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内

 【氏名】 真矢 譲

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社 日立製作所

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ホットスタンバイ計算機システム

【特許請求の範囲】

【請求項 1】 運用系切替により、運用系と待機系との運用が可能な複数の計算機と、該計算機によりアクセスされるデータを格納する共有ディスク装置とを有する計算機システムであって、

前記計算機は、

記憶部と、制御部とを有し、

前記制御部は、

システムを起動したときに、前記共有ディスク装置内の構成情報を取得し、該構成情報に基づいて、該共有ディスク装置に対するアクセス要求を送信することができる活動状態に設定しておき、該共有ディスク装置に対するアクセス要求を受信したときに、該共有ディスク装置にアクセス要求を送信するドライバ手段と

アクセス先ごとに禁止するアクセス要求を示した管理テーブルに基づいて、アプリケーションプログラムを実行することにより発行されたアクセス要求を送信するかどうかを判定し、禁止されていないアクセス先へのアクセス要求を前記ドライバ手段に送信するアクセス制御手段とを実行する計算機システム。

【請求項 2】 請求項 1 に記載の計算機システムであって、

前記アクセス制御手段は、

該計算機が待機系で運用している場合であって運用系計算機に障害が発生したときは、実行するアプリケーションプログラムによるアクセス先に対するアクセス要求を禁止するように前記管理テーブルに登録することを特徴とする計算機システム。

【請求項 3】 請求項 1 に記載の計算機システムであって、

さらに、オペレータにより入力された運用系切替コマンドを前記計算機に対して送信するコンソールとを有し、

前記アクセス制御手段は、

該計算機が待機系で運用している場合であって前記系切替コマンドを受信した

ときは、実行するアプリケーションプログラムによるアクセス先に対するアクセス要求を禁止するように前記管理テーブルに登録することを特徴とする計算機システム。

【請求項 4】 請求項 2 又は 3 に記載の計算機システムであって、

前記アクセス要求として、少なくともライトを禁止するように前記管理テーブルに登録することを特徴とする計算機システム。

【請求項 5】 請求項 1 の計算機システムであって、

前記管理テーブルには、アクセス先ごとの禁止するリード又はライトのアクセス要求を示しており、

前記アクセス制御手段は、前記管理テーブルに基づいて、アプリケーションプログラムを実行することにより発行されたリード又はライトのアクセス要求を送信するかどうかを判定し、リード又はライトが禁止されていないアクセス先へのアクセス要求であるときはリード又はライトのアクセス要求を前記ドライバ手段に送信することを特徴とする計算機システム。

【請求項 6】

請求項 1 の計算機システムであって、

前記管理テーブルには、アクセス先ごとの禁止するファイルオープン又はファイルクローズのアクセス要求を示しており、

前記アクセス制御手段は、前記管理テーブルに基づいて、アプリケーションプログラムを実行することにより発行されたファイルオープン又はファイルクローズのアクセス要求を送信するかどうかを判定し、ファイルオープン又はファイルクローズが禁止されていないアクセス先へのアクセス要求であるときはファイルオープン又はファイルクローズのアクセス要求を前記ドライバ手段に送信することを特徴とする計算機システム。

【請求項 7】 請求項 1 に記載の計算機システムであって、

さらに、オペレータにより入力された、アクセス先ごとに禁止するアクセス要求を登録、削除又は変更するためのコマンドを前記計算機に対して送信するコンソールとを有し、

前記アクセス制御手段は、

前記コマンドを受信したときに、アクセス先IDと、そのアクセス先に対して禁止するアクセス要求の種類とを、前記管理テーブルに登録、削除又は変更することを特徴とする計算機システム。

【請求項 8】 請求項1に記載の計算機システムであって、

さらに、オペレータにより入力された管理テーブルの内容を要求するコマンドを前記計算機に対して送信し、該計算機から受信した管理テーブルの内容を出力するコンソールとを有することを特徴とする計算機システム。

【請求項 9】

運用系切替により、運用系と待機系との運用が可能な計算機であって、
記憶部と、制御部とを有し、
前記制御部は、

システムを起動したときに、複数の計算機によりデータを共有される共有ディスク装置内の構成情報を取得し、該構成情報に基づいて、該共有ディスク装置に対するアクセス要求を送信することができる活動状態に設定しておき、該共有ディスク装置に対するアクセス要求を受信したときに、該共有ディスク装置にアクセス要求を送信するドライバ手段と、

アクセス先ごとに禁止するアクセス要求を示した管理テーブルに基づいて、アプリケーションプログラムを実行することにより発行されたアクセス要求を送信するかどうかを判定し、禁止されていないアクセス先へのアクセス要求を前記ドライバ手段に送信するアクセス制御手段とを実行する計算機。

【請求項 10】 計算機を、運用系切替により運用系と待機系との運用ができるように機能させるためのプログラムであって、

システムを起動したときに、複数の計算機によりデータを共有される共有ディスク装置内の構成情報を取得し、該構成情報に基づいて、該共有ディスク装置に対するアクセス要求を送信することができる活動状態に設定するようにドライバ手段に指示する指示手段と、

アクセス先ごとに禁止するアクセス要求を示した管理テーブルに基づいて、アプリケーションプログラムを実行することにより発行されたアクセス要求を送信するかどうかを判定し、禁止されていないアクセス先へのアクセス要求を前記ド

ライバ手段に送信する判定手段として前記計算機を機能させるためのプログラム。

【請求項 1 1】 計算機を、運用系切替により運用系と待機系との運用ができるように機能させるためのプログラムであって、

システムを起動したときに、複数の計算機によりデータを共有される共有ディスク装置内の構成情報を取得し、該構成情報に基づいて、該共有ディスク装置に対するアクセス要求を送信することができる活動状態に設定するようにしておき、該共有ディスク装置に対するアクセス要求を受信したときに、該共有ディスク装置にアクセス要求を送信する送信手段と、

アクセス先ごとに禁止するアクセス要求を示した管理テーブルに基づいて、アプリケーションプログラムを実行することにより発行されたアクセス要求を送信するかどうかを判定し、禁止されていないアクセス先へのアクセス要求を前記送信手段に送信する判定手段として前記計算機を機能させるためのプログラム。

【請求項 1 2】 請求項 1 1 のプログラムであって、

さらに、OSとして機能することを特徴とするプログラム。

【請求項 1 3】 運用系切替により運用系と待機系との運用可能な計算機によるアクセス制御方法であって、

システムを起動したときに、複数の計算機によりデータを共有される共有ディスク装置内の構成情報を取得し、該構成情報に基づいて、該共有ディスク装置に対するアクセス要求を送信することができる活動状態に設定しておき、該共有ディスク装置に対するアクセス要求を受信したときに、該共有ディスク装置にアクセス要求をドライバ手段により送信し、

アクセス先ごとに禁止するアクセス要求を示した管理テーブルに基づいて、アプリケーションプログラムを実行することにより発行されたアクセス要求を送信するかどうかを判定し、禁止されていないアクセス先へのアクセス要求を前記ドライバ手段に送信するアクセス制御方法。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、現用系と待機系およびその共有ディスクから構成されるホットスタンバイシステムに係わり、共有ディスクの高速切替え方式に関する。

【0002】

【従来の技術】

オンラインランザクションの分野では、可用性向上のため、現用系計算機と、待機系計算機と、共有ディスク装置とから構成されるホットスタンバイのシステム構成が適用されている。

【0003】

従来のシステム構成のホットスタンバイ切替手順を以下に説明する。

【0004】

まず、現用系計算機の障害を検出すると、待機系計算機は、共有ディスク装置へのアクセス要求の発行が可能となるように活動化コマンドをデバイスドライバ手段に発行する。次に、デバイスドライバ手段は、ディスク装置にディスク内の構成情報と性能情報とを要求し、受付けたディスク内の構成情報等に基づいて、論理ボリュームごとに物理ディスク内の領域を特定するマップを作成することで、共有ディスクを確保する。

【0005】

尚、従来のシステム構成のホットスタンバイ切替技術としては、例えば、特許文献1がある。

【0006】

【特許文献1】

特開平10-289122号公報、要約

【0007】

【発明が解決しようとする課題】

従来のホットスタンバイ技術では、計算機は、障害を検出した後に、共有ディスク装置からディスク内の構成情報等を受付けて、その情報に基づいて、ディスクの確保を実施していた。

【0008】

従って、ホットスタンバイ切替処理に長時間かかっていたため、アプリケーション

ョンプログラムの実行待機状態・処理遅延状態が生じ、さらに、障害計算機によるアクセスが生じることによりデータ破壊等が生じる可能性があった。

【0009】

本発明の目的は、高速にホットスタンバイ切替処理を実行する技術を提供することにある。

【0010】

【課題を解決するための手段】

本発明の一実施形態の計算機システムでは、運用系切替により、運用系と待機系との運用が可能な複数の計算機と、計算機によりアクセスされるデータを格納する共有ディスク装置とを有する。この計算機は、記憶部と、ドライバ手段とアクセス制御手段とを実行する制御部とを有する。このドライバ手段は、システムを起動したときに、共有ディスク装置内の構成情報を取得し、構成情報に基づいて、共有ディスク装置に対するアクセス要求を送信することができる活動状態に設定しておく。また、共有ディスク装置に対するアクセス要求を受信したときに、共有ディスク装置にアクセス要求を送信する。また、アクセス制御手段は、アクセス先ごとに禁止するアクセス要求を示した管理テーブルに基づいて、アプリケーションプログラムを実行することにより発行されたアクセス要求を送信するかどうかを判定し、禁止されていないアクセス先へのアクセス要求をドライバ手段に送信する。

【0011】

また、上述した実施形態では、アクセス制御手段は、計算機が待機系で運用している場合であって運用系計算機に障害が発生したときは、実行するアプリケーションプログラムによるアクセス先に対するアクセス要求を禁止するように管理テーブルに登録するようにすることが好ましい。

【0012】

また、上述した実施形態では、さらに、オペレータにより入力された運用系切替コマンドを計算機に対して送信するコンソールとをすることが好ましい。また、アクセス制御手段は、計算機が待機系で運用している場合であって系切替コマンドを受信したときは、実行するアプリケーションプログラムによるアクセス先

に対するアクセス要求を禁止するように管理テーブルに登録することが好ましい。

【0013】

さらに、アクセス制御手段は、前記アクセス要求として、少なくともライトを禁止するように前記管理テーブルに登録することが好ましい。

【0014】

また、上述した実施形態では、管理テーブルには、アクセス先ごとの禁止するリード又はライトのアクセス要求を示していることが好ましい。また、アクセス制御手段は、管理テーブルに基づいて、アプリケーションプログラムを実行することにより発行されたリード又はライトのアクセス要求を送信するかどうかを判定し、リード又はライトが禁止されていないアクセス先へのアクセス要求であるときはリード又はライトのアクセス要求をドライバ手段に送信することが好ましい。

【0015】

また、上述した実施形態では、管理テーブルには、アクセス先ごとの禁止するファイルオープン又はファイルクローズのアクセス要求を示していることが好ましい。さらに、アクセス制御手段は、管理テーブルに基づいて、アプリケーションプログラムを実行することにより発行されたファイルオープン又はファイルクローズのアクセス要求を送信するかどうかを判定し、ファイルオープン又はファイルクローズが禁止されていないアクセス先へのアクセス要求であるときはファイルオープン又はファイルクローズのアクセス要求を前記ドライバ手段に送信することが好ましい。

【0016】

また、上述した実施形態では、さらに、オペレータにより入力された、アクセス先ごとに禁止するアクセス要求を登録、削除又は変更するためのコマンドを前記計算機に対して送信するコンソールとを有することが好ましい。さらに、アクセス制御手段は、コマンドを受信したときに、アクセス先IDと、そのアクセス先に対して禁止するアクセス要求の種類とを、管理テーブルに登録、削除又は変更することが好ましい。

【 0 0 1 7 】

また、上述した実施形態では、さらに、オペレータにより入力された管理テーブルの内容を要求するコマンドを前記計算機に対して送信し、計算機から受信した管理テーブルの内容を出力するコンソールとを有することが好ましい。

【 0 0 1 8 】

【発明の実施の形態】

図1は、本発明の一実施形態である計算機システムの構成図である。

【 0 0 1 9 】

この計算機システムは、現用系計算機(1000)と、待機系計算機(2000)と、共有ディスク装置(3000)と、コンソール(4000)とを有する。

【 0 0 2 0 】

現用系計算機(1000)と待機系計算機(2000)は、それぞれ、プロセッサ(1500、2500)と、メモリ(1600、2600)と、SVP(Service Processor)(1700、2700)と、入出力インタフェース(以下、I/F)(1800、2800)と、主記憶装置(1650、2650)と、系切替制御手段(1100、2100)と、アクセス制御手段(1200、2200)と、OS(オペレーティングシステム)(1300、2300)と、デバイスドライバ手段(1400、2400)とを有する。

【 0 0 2 1 】

共有ディスク装置(3000)は、現用系計算機(1000)及び待機系計算機(2000)のアプリケーションプログラム(1410、2410)を実行することによりアクセスされるデータを、一又は複数のボリューム単位で格納している。尚、ここで、ボリューム単位とは、物理ディスク等の物理的な単位であっても、ファイル等の論理的な単位であってもよい。

【 0 0 2 2 】

コンソール(4000)は、SVP(1700、2700)に接続されており、オペレータにより入力されたコマンドを受付け、受付けたコマンドに応じて、計算機(1000、2000)内のいずれかの手段に対して、コマンドを発行する。

【 0 0 2 3 】

ここで、系切替制御手段(1100、2100)と、アクセス制御手段(1200、2200)と、

OS(1300、2300)と、デバイスドライバ手段(1400、2400)として機能させる為のプログラムは、ROM等の記録媒体に記録され、主記憶装置(1650)に格納された後、メモリ(1600、2600)にロードされてプロセッサ(1500、2500)により実行されるものとする。プログラムを記録する媒体は、CD-ROM以外の他の記憶媒体でも良い。また、プログラムは、当該記憶媒体からメモリ(1600、2600)にインストールしても良いし、ネットワークを通じて当該記憶媒体にアクセスしてプログラムを使用するものとしても良い。尚、系切替制御手段(1100、2100)と、アクセス制御手段(1200、2200)と、OS(1300、2300)と、デバイスドライバ手段(1400、2400)として機能させる為のハードウェア構成を現用系計算機(1000)又は待機系計算機(2000)のプロセッサ(1500、2500)とは独立して備えるようにしてもよい。

【0024】

系切替制御手段(1100、2100)は、現用系計算機(1000)あるいは待機系計算機(2000)の障害を検出し、後述するホットスタンバイ切替処理をするように指示する。尚、系状態としては、ホットスタンバイシステムでは、以下の大きくは以下の3状態に分類される。実際にサービスを提供している現用系状態と、サービスは提供していないが、現用系で障害が発生すると、直ちに処理を引き継げる状態待機状態と、障害状態であるオフライン状態とである。

【0025】

系切替制御手段(1100、2100)は、各々一定間隔でaliveメッセージを送信し、もう片方の系がそれを受信する。現用系計算機(1000)で障害が発生するとaliveメッセージが送信されなくなる。このとき、待機系の系切替制御手段(1100、2100)は、現用系計算機(1000)で障害が発生したことを検出し、後述するホットスタンバイ切替制御を実行する。

【0026】

アクセス制御手段(1200、2200)は、アプリケーションプログラム(1410、2410)を実行することによりアクセスされる特定のボリュームに対するアクセス要求の態様を、管理テーブル(1295、2295)に従って制御する。ここで、アクセス要求とは、例えば、リード要求、ライト要求、ファイルオープン要求やファイルクロー

ズ要求がある。また、アクセス要求の態様とは、アクセス要求をデバイスドライバ手段(1400、2400)に発行することを許可するか禁止するかの態様をいい、例えば、リードとライトともに許可又は（以下、／と示す。）禁止、リードは許可／禁止であるがライトは禁止／許可するような態様が考えられる。尚、共有ディスク装置(3000)を構成するボリュームをグループ化し、各グループに固有のグループIDを割当てて、グループ単位でアクセス要求を制限してもよい。

【 0 0 2 7 】

OS(1300、2300)は、アプリケーションプログラム(1410、2410)から発行されるアクセス要求と図8に示すデバイススイッチテーブルとに基づいて、アクセス要求の発行先を特定する。デバイススイッチテーブルとは、アクセス要求先のボリュームIDと、アクセス要求の態様との組合せごとに、アクセス要求の発行先を示すテーブルである。アクセス要求の発行先には、もともとデバイスドライバ手段(1400、2400)のアドレス（図8中ドライバID）が示されているが、後述するように、アクセス態様登録手段(1210、2210)により登録されたボリュームIDとアクセス要求の態様との組合せには、アクセス制御手段(1200、2200)のアドレス（図8中アクセスID）が示されている。

【 0 0 2 8 】

また、OS(1300、2300)は、活動状態／非活動状態となるように活動化コマンド／非活動化コマンドをデバイスドライバ手段(1400、2400)に発行する。尚、コンソールは、オペレータにより入力された活動化コマンド／非活動化コマンドを受け付けることができる。また、活動状態／非活動状態についての詳細は、図4を用いて説明する。

【 0 0 2 9 】

デバイスドライバ手段(1400、2400)は、ディスク内の構成情報と、ディスク内の構成要素ごとの性能情報を取得し、アクセス要求をディスク内のどの物理領域に割当てて示す変換マップを作成し、保持する。

【 0 0 3 0 】

デバイスドライバ手段(1400、2400)は、アクセス要求をOS(1300、2300)又はアクセス制御手段(1200、2200)から受信すると、アクセス要求と、ディスク内の

構成情報と、ディスク内の構成要素ごとの性能情報とに基づいて、アクセス要求をディスク内の物理領域を割当て、共有ディスク装置(3000)に発行する。

【 0 0 3 1 】

図2は、アクセス制御手段(1200、2200)の構成を示す図である。

【 0 0 3 2 】

アクセス制御手段(1200、2200)は、アクセス態様登録手段(1210、2210)と、アクセス態様取得手段(1270、2270)、アクセス態様判定手段(1290、2290)と、管理テーブル(1295、2295)とからなる。

【 0 0 3 3 】

アクセス態様登録手段(1210、2210)は、管理テーブル(1295、2295)に、ボリュームのIDを登録・削除し、登録したボリュームIDごとのアクセス要求の態様を登録・変更する。尚、アクセス態様登録手段(1210、2210)は、共有ディスク装置(3000)にボリュームが新たに定義された場合や物理ディスクが追加された場合等は、ボリュームIDとそのアクセス要求の態様を、コンソールを介して受け付けることができる。また、複数のボリュームにグループIDを指定することにより、グループ単位でボリュームを登録・削除し、グループ単位でアクセス要求の態様の登録・変更をすることもできる。

【 0 0 3 4 】

アクセス態様取得手段(1270、2270)は、コンソール(4000)又はプロセッサ(1500、2500)からの問合せに応じ、指定されたボリューム又はグループのアクセス要求の態様をコンソール(4000)に返す。コンソール(4000)は、受信したボリューム又はグループのアクセス要求の態様を出力する。

【 0 0 3 5 】

アクセス態様判定手段(1290、2290)は、OS(1300、2300)からのアクセス要求と管理テーブル(1295、2295)とに基づいて、デバイスドライバ手段(1400、2400)にアクセス要求を発行するかどうかを判定する。例えば、アクセス要求がリードを許可／禁止しているボリュームに対してである場合、デバイスドライバ手段(1400、2400)からのアクセスを許可／禁止する。また、アクセス要求がライトを許可／禁止しているボリュームに対してである場合、デバイスドライバ手段(1400、2

400)へのアクセスを許可／禁止する。アクセス要求がいずれのアクセス態様も許可していないボリュームに対してである場合、デバイスドライバ手段(1400、2400)との間のアクセスを禁止する。

【 0 0 3 6 】

管理テーブル(1295、2295)は、図9に示すように、ボリュームIDとアクセス要求との組合せごとのアクセス要求の態様を示すテーブルである。尚、複数のボリュームにグループIDを指定することにより、グループ単位でアクセス要求の態様を記録することもできる。管理テーブルへは、ボリュームIDとそのアクセス要求との組合せごとに登録することができる。例えば、図9では、V o 1 3のボリュームIDにおいては、ファイルオープンのアクセス要求は登録されているが、リード及びライト要求は登録されていない(図9中「-」で示してある。)。また、図8に示すデバイススイッチテーブルにはV o 1 4のボリュームIDがあるが、図9の管理テーブルには登録されていない。尚、図9では、アクセス要求の態様として、許可する場合と禁止する場合の双方を登録してあるが、禁止する場合のみ登録するようにしてもよい。

【 0 0 3 7 】

図3は、共有ディスク装置(3000)の構成図である。

【 0 0 3 8 】

VG00(3101)、VG01(3102)、VG02(3103)、VG03(3201)、VG04(3202)、VG05(3301)、VG06(3302)の7つのボリュームを3つのグループに分類し、各グループに0～2のIDを割り当て、それぞれアクセス要求の態様を設定した例を示している。尚、ボリューム単位でIDを割り当ててもよい。

【 0 0 3 9 】

グループID=0のグループ(3100)はリードおよびライトを許可しており、リード、ライト共に通常通りに行われる。

【 0 0 4 0 】

また、グループID=1のグループ(3200)は、ライトのみ禁止しており、リード処理は通常通り行われ、ライトは失敗する。

【 0 0 4 1 】

グループID=2のグループ(3300)はリードとライト共に禁止されており、リード／ライト処理は失敗する。

【 0 0 4 2 】

図4は、ボリュームの状態遷移を示す図である。

【 0 0 4 3 】

ボリュームの状態は、管理対象外状態(5100)と管理対象内状態(5200)に分けられる。

【 0 0 4 4 】

管理対象外状態(5100)は、さらに、活性状態と非活性状態とに分類される。

【 0 0 4 5 】

ここで、活性状態と非活性状態との遷移手順を示す。

【 0 0 4 6 】

まず、非活性状態から活性状態への遷移手順を示す。

【 0 0 4 7 】

非活動状態で、デバイスドライバ手段(1400、2400)は、OS(1300、2300)から活動化コマンドを受信すると、共有ディスク装置(3000)からディスク内の構成情報と性能情報とを要求する。受信したデバイスドライバ手段(1400、2400)は、ディスク内の構成情報と性能情報とに基づいて、ボリュームごとに割当て物理的領域を示す変換マップを作成し、メモリ(1600、2600)に記憶する。これにより、デバイスドライバ手段(1400、2400)は、アクセス要求を受けた場合に、作成した変換マップに基づいて、共有ディスク装置(3000)にアクセス要求を発行することができる。この状態を活性状態という。

【 0 0 4 8 】

次に、活性状態から非活性状態への遷移手順を示す。

【 0 0 4 9 】

活動状態で、デバイスドライバ手段(1400、2400)は、OS(1300、2300)から非活動化コマンドを受信すると、メモリ(1600、2600)に記憶された変換マップを破棄する。これにより、デバイスドライバ手段(1400、2400)は、アクセス要求を受けた場合であっても、共有ディスク装置(3000)にアクセス要求を発行することが

できない。この状態を非活性状態という。

【 0 0 5 0 】

ここで、管理対象外状態(5100)と管理対象内状態(5200)との間の遷移手順を示す。

【 0 0 5 1 】

まず、管理対象外状態(5100)から管理対象内状態(5200)に遷移させる手順を示す。アクセス態様登録手段(1210、2210)は、コンソールから指定されたボリュームIDとそのアクセス要求の態様を、管理テーブル(1295、2295)に登録する。また、アクセス態様登録手段(1210、2210)は、管理テーブル(1295、2295)に登録したボリュームIDとそのアクセス要求の態様に該当するアクセス要求先アドレスにアクセス態様登録手段(1210、2210)のアドレスを書込む。これにより、OS(1300、2300)は、受信したアクセス要求が、管理テーブル(1295、2295)に登録したボリュームでかつ登録したアクセス要求の態様である場合には、受信したアクセス要求をアクセス態様登録手段(1210、2210)に発行する。また、アクセス態様登録手段(1210、2210)は、管理テーブル(1295、2295)に従って、受信したアクセス要求をデバイスドライバ手段に発行するかどうかを制御することができる。

【 0 0 5 2 】

このように、アクセス態様登録手段(1210、2210)は、管理テーブル(1295、2295)に従って、受信したアクセス要求をデバイスドライバ手段(1400、2400)に発行するかどうかを制御することができる状態を管理対象内状態(5200)という。逆に、管理テーブル(1295、2295)から登録されたボリュームを除外することにより、管理対象内状態(5200)から管理対象外状態(5100)へ遷移させることができる。

【 0 0 5 3 】

管理対象内状態(5200)は、本実施例では、リードとライトともに許可の態様(5210)、リードは許可／禁止でライトは禁止／許可の態様(5220)、リードとライトともに禁止の態様(5230)に分類される。

【 0 0 5 4 】

図5は、アプリケーションプログラムを実行することにより生じるアクセス処理を模式的に説明する為の図である。

【 0 0 5 5 】

アプリケーションプログラム(1410)は、実行中にOS(1300)に対してリード、ライト等のアクセス要求を発行する(処理5610)。このとき、アクセス要求には、少なくともボリュームIDと、リード／ライト等のアクセス要求の種類とを示す情報が含まれている。

【 0 0 5 6 】

OS(1300)は、アプリケーションプログラム(1410)からの要求に応じ、デバイススイッチテーブル(1350)に基づいて、アクセス先を決定し、アクセス要求を発行する。

ここで、管理テーブル(1295、2295)に登録されたボリュームに対するアクセス要求である場合は、デバイススイッチテーブル(1350)にアクセス制御手段(1200、2200)のアドレスが登録されているので、OS(1300)は、受信したアクセス要求をアクセス態様登録手段(1210)に発行する(処理5620)。

【 0 0 5 7 】

アクセス態様判定手段(1290)は、管理テーブル(1295、2295)に基づいて、デバイスドライバ手段(1400、2400)にアクセス要求を発行するかどうかを判定する。管理テーブル(1295、2295)にアクセス許可として登録されているボリュームに対するアクセス要求であれば、許可されたアクセス要求の態様で、デバイスドライバ手段(1400)にアクセス要求を発行する(処理5630)。ここで、「許可されたアクセス要求の態様で、」とは、例えば、アクセス要求態様がリードの許可／禁止である場合は、デバイスドライバ手段(1400)へのリードを許可／禁止し、アクセス要求態様がライトの許可／禁止である場合は、デバイスドライバ手段(1400)へのライトを許可／禁止すること意味する。

【 0 0 5 8 】

デバイスドライバ手段(1400)は、変換マップに基づいて、アクセス態様判定手段(1290)から要求されたボリュームへのアクセス要求を発行する。

【 0 0 5 9 】

アクセス態様判定手段(1290)は、管理テーブル(1295)にアクセス禁止として登録されているボリュームに対するアクセス要求を受信した場合は、エラーで終了

する(処理5640)。尚、アクセス態様判定手段(1290)は、このエラーの通知をコンソールに対して出力するように指示し、コンソールは、エラーを出力するようにしてもよい。

【0060】

図6は、ホットスタンバイ処理の処理内容を示す図である。

【0061】

現用系計算機(1000)のアクセス制御手段(1200)は、システム立ち上げ後に、OS(1300)に活動化コマンドを発行するように指示する。現用系計算機(1000)のOS(1300)は、活動コマンド(5101)をデバイスドライバ手段(1400)に発行する。デバイスドライバ手段(1400)は、共有ディスク装置(3000)を構成するボリュームを活動状態(5120)に設定する(ステップ7000)。

【0062】

次に、アクセス態様登録手段(1210)は、初期値に基づいて、ボリュームIDを管理テーブル(1295)に登録する(ステップ7050)。ここで、初期値とは、予めオペレータにより設定された管理テーブル(1295、2295)の内容を示す値である。

【0063】

更に、アクセス態様登録手段(1210)は、初期値に基づいて、登録したボリュームIDごとのアクセス要求の態様を管理テーブル(1295)に登録する(ステップ7100)。

【0064】

一方、待機系計算機(2000)のアクセス制御手段(2200)は、システム立ち上げ後に、OS(2300)に活動化コマンドを発行するように指示する。待機系計算機(2000)のOS(2300)は、活動コマンド(5101)をデバイスドライバ手段(2400)に発行する。デバイスドライバ手段(2400)は、共有ディスク装置(3000)を構成するボリュームを活動状態(5120)に設定する(ステップ7500)。

【0065】

次に、アクセス態様登録手段(2210)は、初期値に基づいて、特定のボリュームIDを管理テーブル(2295)に登録する。(ステップ7550)

更に、アクセス態様登録手段(2210)は、登録されたボリュームIDのアクセス

態様をリード及びライトを禁止として、管理テーブル(2295)に登録する(ステップ7600)。尚、アクセス態様として、リードを許可し、ライトを禁止するように登録してもよい。これにより、待機系計算機(2000)において、アプリケーションプログラムによるリード要求を許可することができる。

【 0 0 6 6 】

ステップ7000、7050、7100により、運用系計算機(1000)のアクセス制御判定手段(1290)は、運用系計算機(1000)のアプリケーションプログラム(1410)を実行することにより発行されるリード／ライトのアクセス要求を、管理テーブル(1295)に基づいて、デバイスドライバ手段(1400)に発行することができる運用系状態となる。

一方、ステップ7500、7550、7600により、待機系計算機(2000)のアクセス制御判定手段(2290)は、待機系計算機(2000)のアプリケーションプログラム(2410)を実行することにより発行されるリード及びライトのアクセス要求を、デバイスドライバ手段(2400)に発行しない待機系状態となる。

【 0 0 6 7 】

ここまでの、システム立ち上げ後に実行される初期動作である。

【 0 0 6 8 】

次に、現用系計算機(1000)で障害が発生した場合の動作を説明する。

【 0 0 6 9 】

この場合、待機系計算機(2000)の系切替制御手段(2100)は、aliveメッセージの途絶により現用系計算機(1000)の障害を検出する(ステップ7650)。

【 0 0 7 0 】

障害検出後、待機系計算機(2000)の系切替制御手段(2100)は、現用系計算機(1000)へリセット要求を発行する(ステップ7700)。

【 0 0 7 1 】

現用系計算機の系切替制御手段(1100)は、待機系計算機(2000)の系切替制御手段(2100)が発行したリセット要求を受信する(ステップ7150)。

【 0 0 7 2 】

リセット要求を受信した現用系計算機の系切替制御手段(1100)は、オフライン

状態となるようにシステムコール(5202)を、アクセス制御登録手段(1210)に発行する。アクセス制御登録手段(1210)は、このシステムコール(5202)を受けて、管理テーブル(1295)に登録されたボリュームIDのアクセス要求の態様をリード及びライトを禁止となるように変更する。これにより、運用系計算機のアクセス制御判定手段(1290)は、運用系計算機(1000)のアプリケーションプログラム(1410)を実行することにより発行されるリード及びライトのアクセス要求を、デバイスドライバ手段(1400)に発行しないオフライン状態に切替る(ステップ7200)。

【 0 0 7 3 】

尚、現用系計算機の系切替制御手段(1100)は、リセット要求を受信したときに、現用系計算機のOS(1300)に非活性化コマンドをデバイスドライバ手段(1400)に発行するようように指示してもよい。この場合、デバイスドライバ手段(1400)は、OS(1300)から非活動化コマンドを受信すると、メモリ(1600)に記憶された変換マップを破棄する。これにより、デバイスドライバ手段(1400)は、アクセス要求を受けた場合であっても、共有ディスク装置(3000)にアクセス要求を発行することができなくなり、オフライン状態となる。さらに、現用系計算機(1000)は、障害から修復すると、システム立ち上げ時と同様にボリュームの活動化コマンド(5101)を発行し、非活動化状態(5110)から活動化状態(5120)に遷移させることにより、オフライン状態となるようにしてもよい。

【 0 0 7 4 】

次に、系切替制御手段(1100)は、運用系計算機(1000)がオフライン状態に切替ったことのリセット完了通知を待機系計算機(2000)に対して送信する(ステップ7300)。

【 0 0 7 5 】

障害の発生した現用系は、障害から修復すると、アクセス態様登録手段(1210)は、初期値に基づいて、登録したボリュームIDごとのアクセス要求の態様を管理テーブル(1295)に登録し、オフライン状態から待機系状態に遷移する(ステップ7400)。

【 0 0 7 6 】

待機系計算機(2000)の系切替制御手段(2100)は、現用系計算機(1000)からのリ

セットを受信すると(ステップ7750)、運用系状態に切替るようにシステムコール(5240)を、アクセス制御登録手段(2210)に発行する。アクセス制御登録手段(2210)は、このシステムコール(5240)を受けて、初期値に基づいて、管理テーブル(2295)を変更する。これにより、待機系計算機(2000)のアクセス状態判定手段(2290)は、管理テーブル(2295)に基づいて、デバイスドライバ手段(2400)に、アクセス要求を発行する運用系状態となる(ステップ7800)。尚、アクセス制御登録手段(2210)は、システムコール(5240)とともに、現用系計算機(1000)から管理テーブル(1295)の内容を受信し、現用系計算機(1000)の管理テーブル(1295)の内容を引継ぐように、管理テーブル(2295)の内容を変更するようにしてもよい。

【 0 0 7 7 】

このようにして、現用系計算機(1000)で障害が発生すると、アクセス態様登録手段(1210、2210)により、管理テーブル(1295、2295)の内容を変更することにより、現用系計算機(1000)はオフライン状態に、待機系計算機(2000)運用系状態になり、共有ディスクの高速切替えを実現する。

【 0 0 7 8 】

図7は、コンソール(4000)から指示によるホットスタンバイ切替えを示す図である。

【 0 0 7 9 】

ここでは、運用系計算機(1000)は、ステップ7000、7050、7100により、既に、初期値に基づいて、アクセス態様登録手段(1210)により、ボリュームのIDと、ボリュームIDごとのアクセス要求の態様が登録されているとする。また、待機系計算機(2000)は、ステップ7500、7550、7600により、システム立ち上げ後の初期動作を実行済みである。

【 0 0 8 0 】

コンソールは、オペレータより、ホットスタンバイ切替えコマンドを受付けると、切替えコマンドを現用系計算機(1000)に発行する(ステップ8000)。

【 0 0 8 1 】

現用系計算機(1000)は、切替えコマンドを受信すると、アクセス態様登録手段(1200)により、管理テーブル(1295)に登録されたボリュームIDのアクセス要求

の態様をリード及びライトを禁止となるように変更する。これにより、運用系計算機のアクセス制御判定手段(1290)は、運用系計算機(1000)のアプリケーションプログラム(1410)を実行することにより発行されるリード及びライトのアクセス要求を、デバイスドライバ手段(1400)に発行しなくなり、運用系計算機(1000)はオフライン状態に切替る(ステップ8100)。尚、現用系計算機(1000)は、切替えコマンドを受信すると、アクセス態様登録手段(1200)により、管理テーブル(1295)に登録されたボリュームID削除するようにしてもよい。これにより、ボリュームを管理対象外状態に遷移させることにより、オフライン状態に切替ることができる。また、現用系計算機(1000)は、切替えコマンドを受信すると、非活動状態に遷移させることにより、オフライン状態に切替るようにしてもよい。

【 0 0 8 2 】

現用系計算機(1000)は、オフライン状態に切替ると、リセット完了通知を、コンソールに送信する(ステップ8200)。

【 0 0 8 3 】

コンソールは、リセット完了通知を受信すると、待機系計算機(2000)に系切替コマンドを発行する(ステップ8300)。

【 0 0 8 4 】

待機系計算機のアクセス状態登録手段(2210)は、系切替コマンドを受信すると、初期値に基づいて、管理テーブル(2295)の内容を書換える(ステップ8400)。尚、系切替コマンドとともに、運用系計算機(1000)で保持していた管理テーブル(1295)の内容を待機系計算機(2000)に送信し、待機系計算機(2000)は、運用系計算機(1000)で保持していた管理テーブル(1295)の内容に基づいて、管理テーブル(2295)の変更を実行するようにしてもよい。

【 0 0 8 5 】

これにより、共有ディスク装置(3000)の系を切替えることができ、待機系計算機(2000)は運用系計算機(1000)の処理を引継ぐことができる。

【 0 0 8 6 】

上述した実施形態では、アクセス制御手段(1200、2200)をデバイスドライバ手段(1400、2400)及びOS(1300、2300)とは独立した構成を採用した場合の実施例に

について説明した。

【 0 0 8 7 】

上記の場合以外にも、アクセス制御手段(1200、2200)をOS(1300、2300)内に配置することも可能である。この場合に、アクセス態様登録手段は、管理テーブルではなく、デバイススイッチテーブルに登録・削除することにより、管理テーブルとデバイススイッチテーブルを統一化してもよい。

【 0 0 8 8 】

また、アクセス制御手段(1200、2200)をデバイスドライバ手段(1400、2400)内に配置することも可能である。この場合に、管理テーブルと上述した変換マップとを統一化してもよい。

【 0 0 8 9 】

尚、上述した実施形態では、アクセス要求の態様として、リード／ライトの許可／禁止について説明したが、アクセス要求の態様として、ファイルのオープン／クローズの許可／禁止についても、同様に制御できる。

【 0 0 9 0 】

【発明の効果】

本発明では、高速にホットスタンバイ切替処理を実行する技術を提供することができる。

【図面の簡単な説明】

【図 1】 本発明の一実施形態である計算機システムの構成図。

【図 2】 アクセス制御手段の構成図。

【図 3】 共有ディスクの構成図。

【図 4】 ボリュームの状態遷移図。

【図 5】 本発明の一実施形態のアクセス要求入出力処理を示す模式図。

【図 6】 本発明の一実施形態の障害時のホットスタンバイ処理の流れ図。

【図 7】 本発明の一実施形態のコンソール入力によるホットスタンバイ処理の流れ図。

【図 8】 デバイススイッチテーブルの内容を示す図。

【図 9】 管理テーブルの内容を示す図。

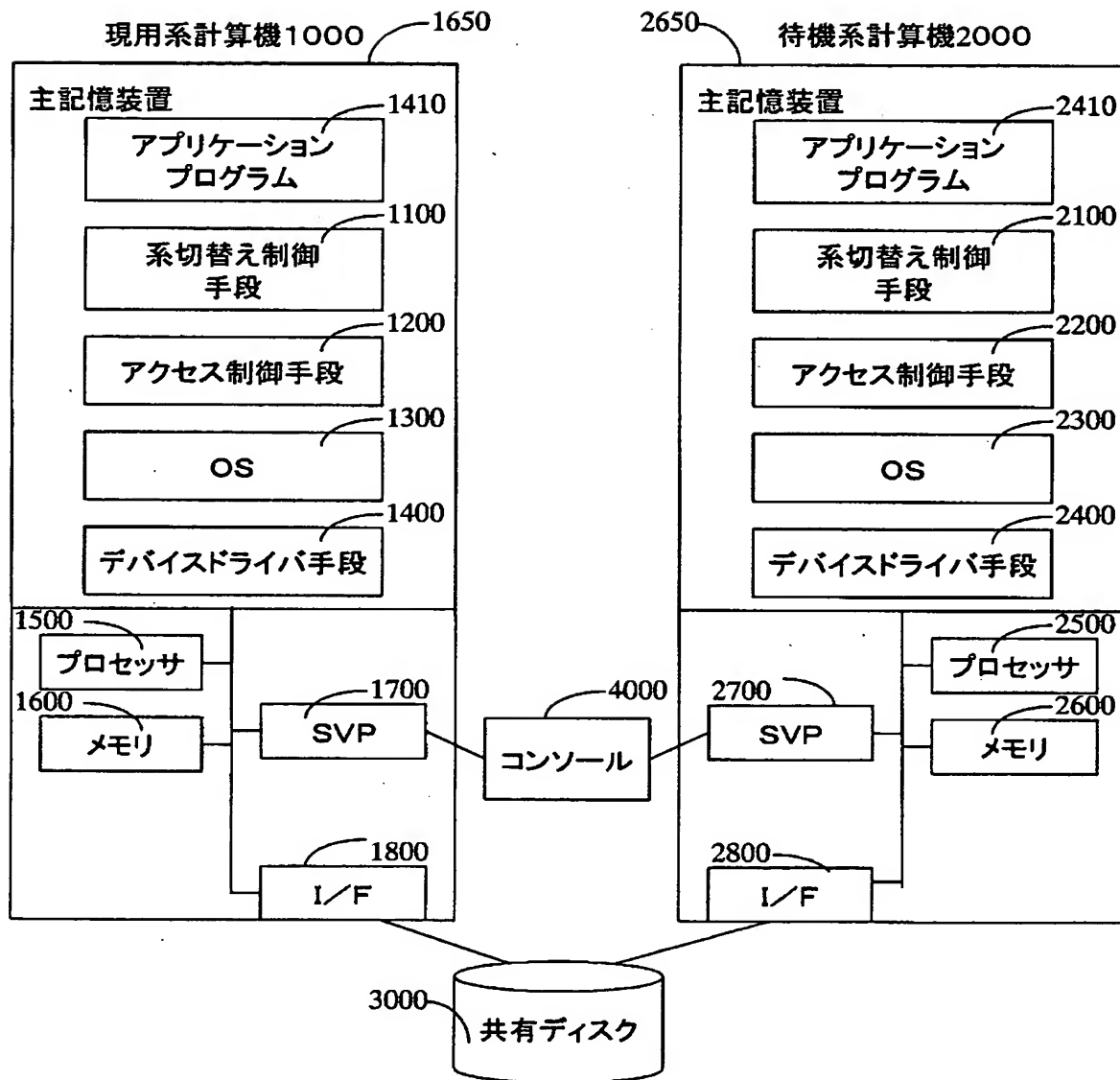
【符号の説明】

1000：現用系、2000：待機系、1100、2100：系切替制御手段、1200、2200：アクセス制御手段、1300、2300：OS、1400、2400：デバイスドライバ手段、1500、2500：プロセッサ、1600、2600：メモリ、1700、2700：SVP(Service Processor)、3000：共有ディスク、4000：コンソール

【書類名】 図面

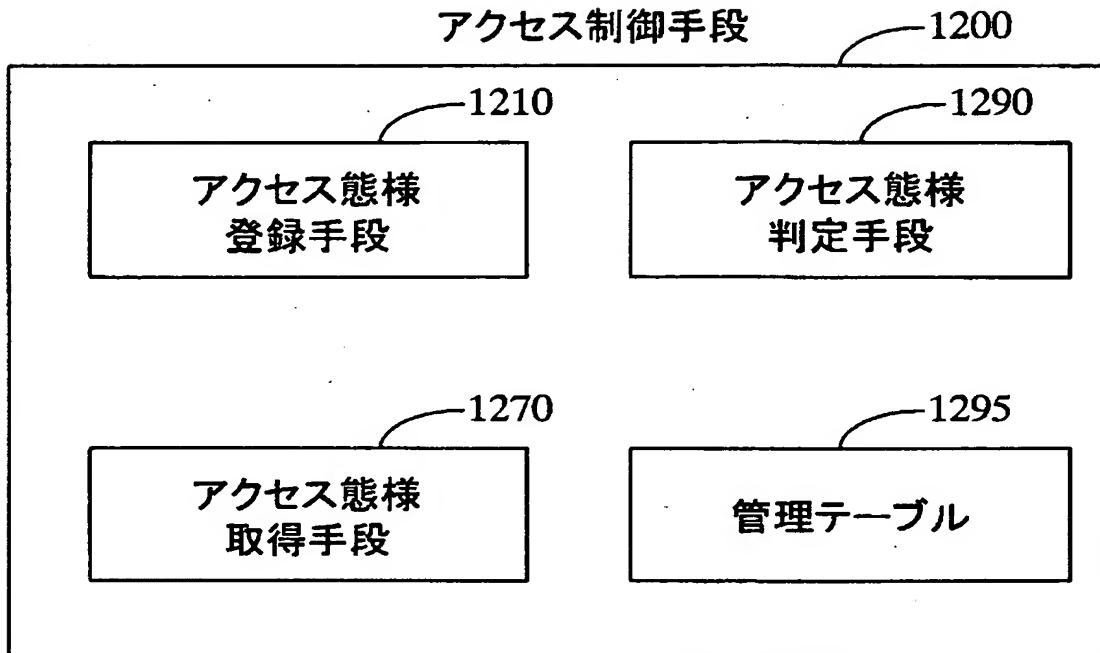
【図 1】

図1



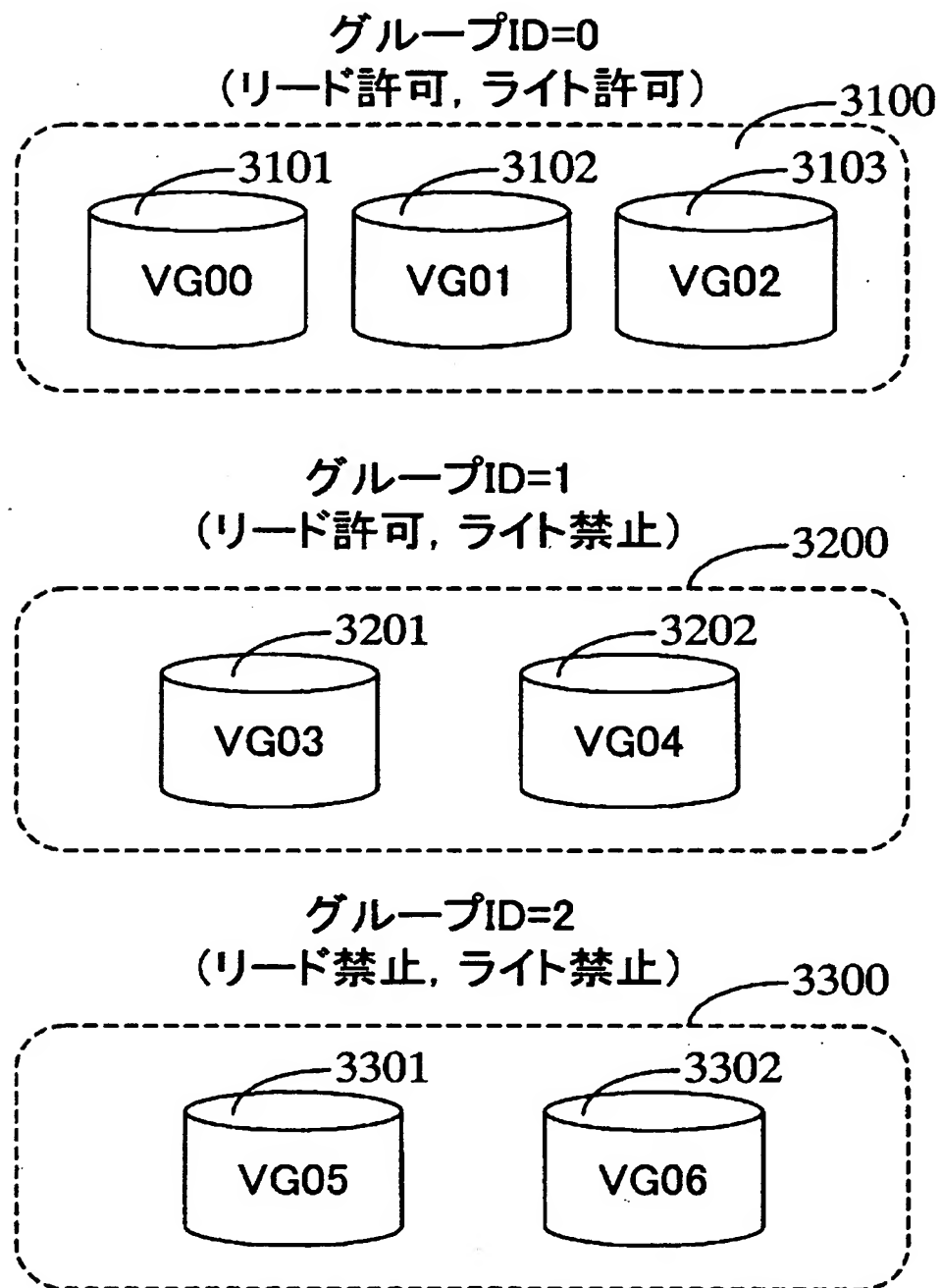
【図 2】

図2

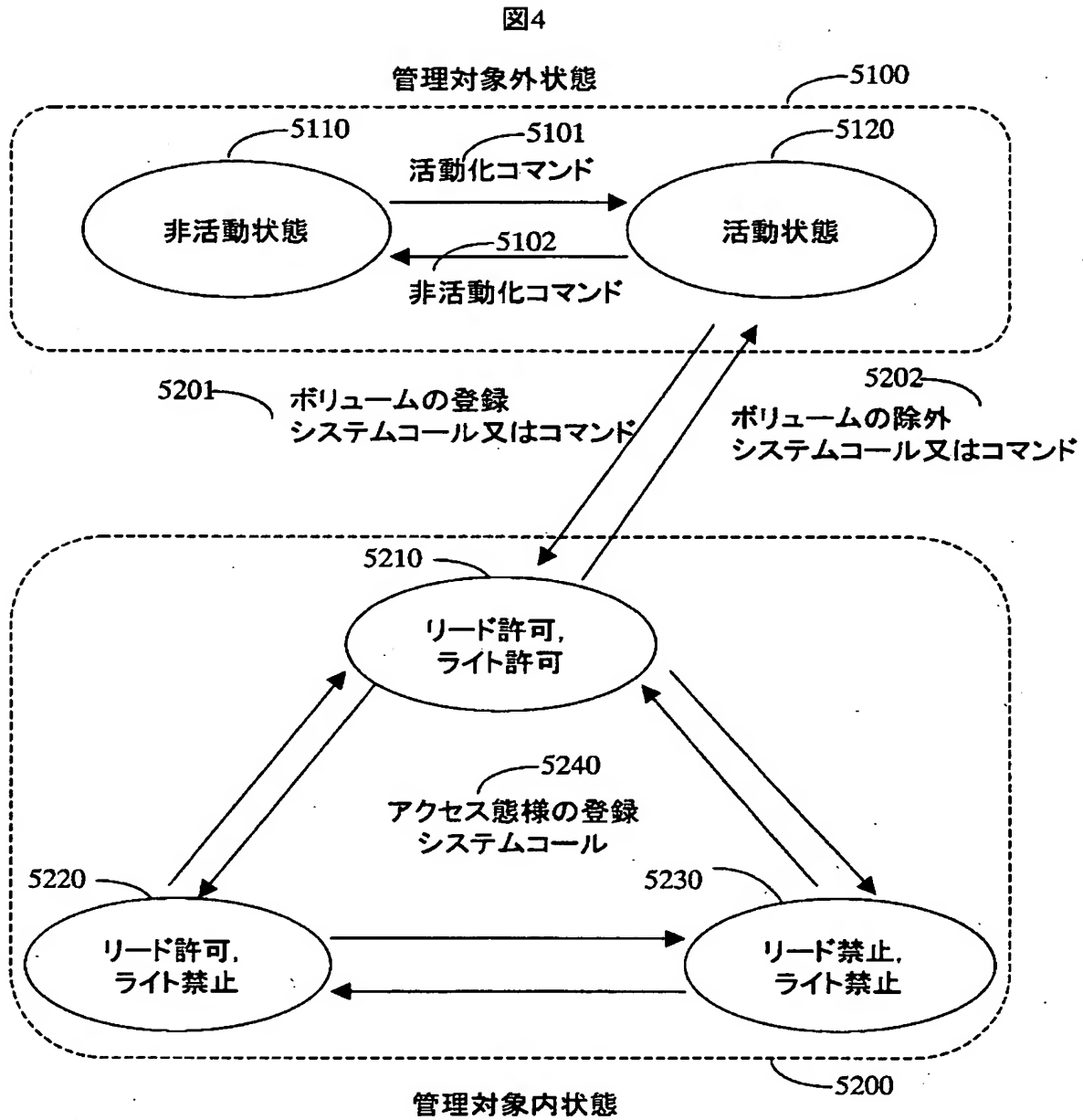


【図 3】

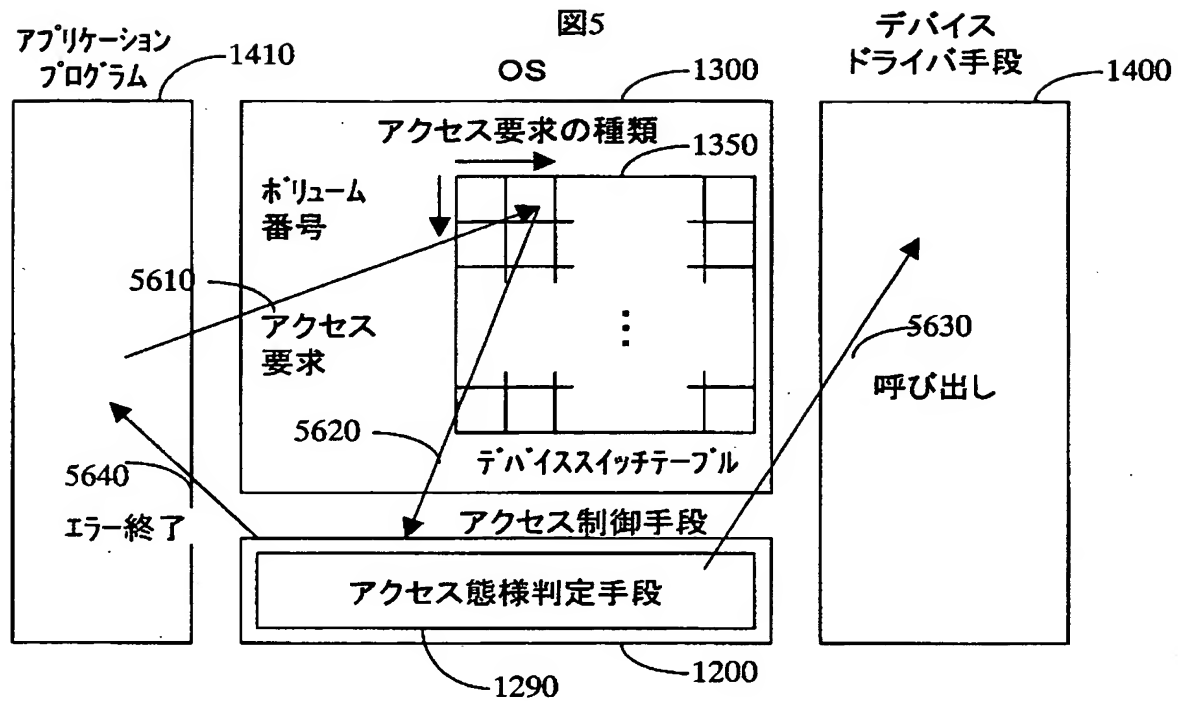
図3



【図 4】

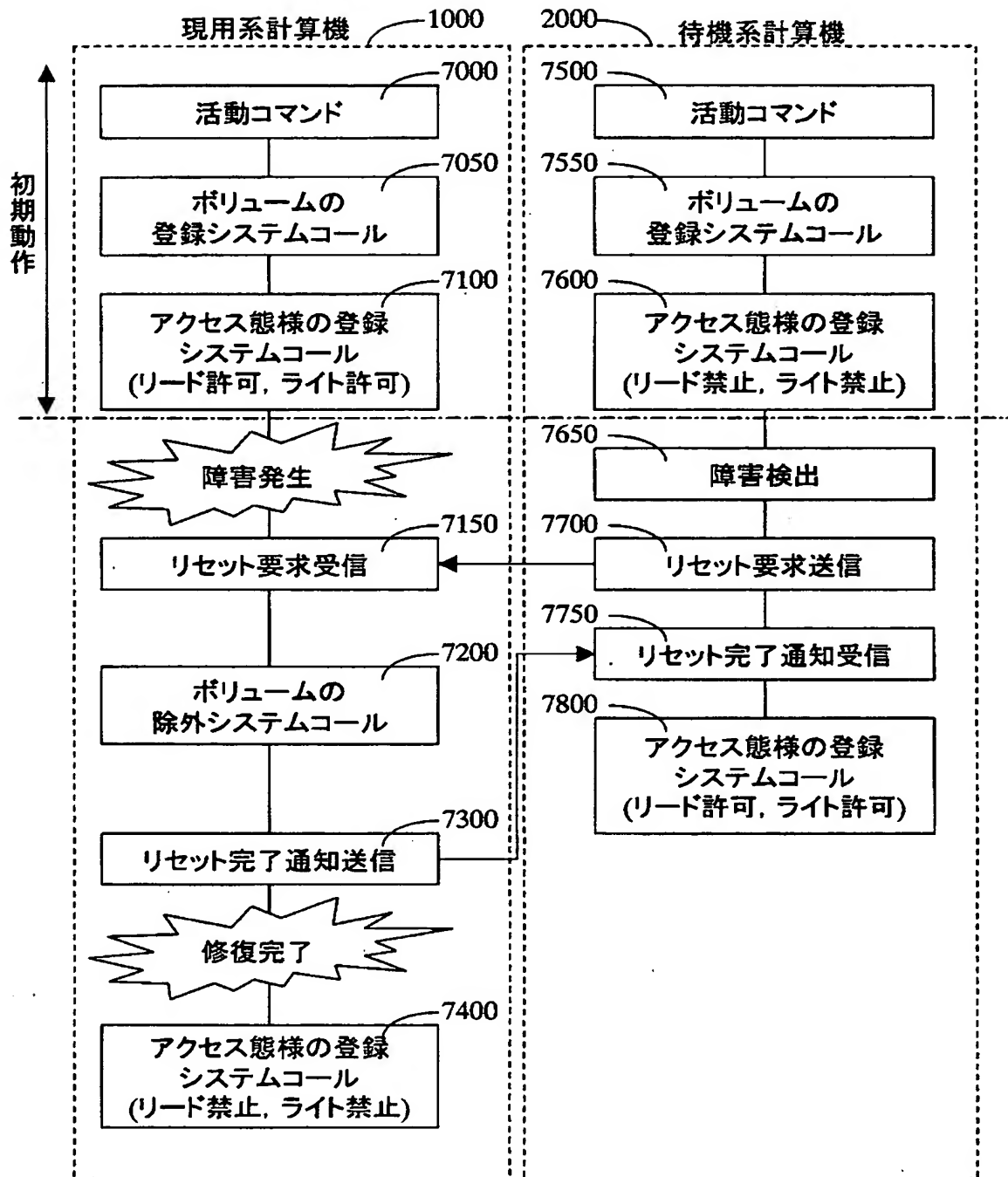


【図5】

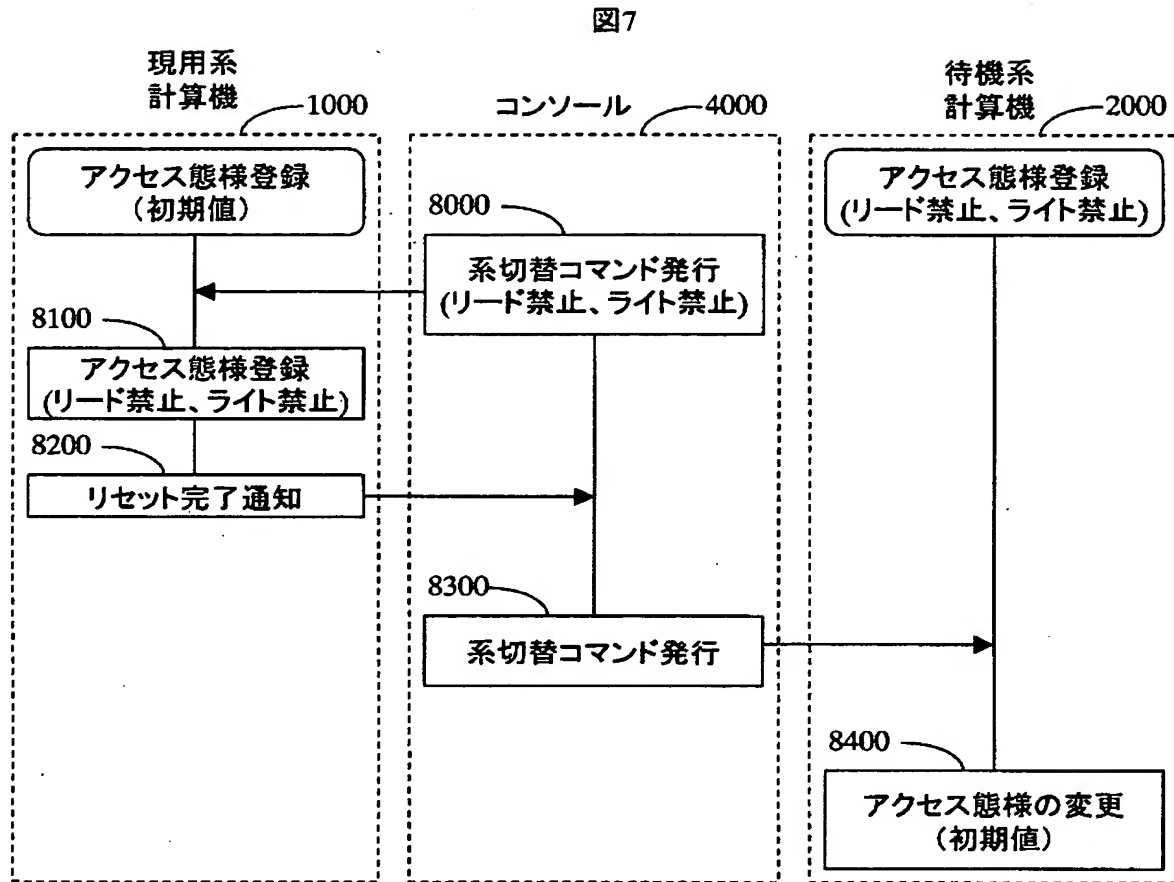


【図 6】

図6



【図 7】



【図 8】

図8

アクセス要求 ボリュームID	リード	ライト	ファイルオープン	...
Vol0	アクセスID	アクセスID	アクセスID	...
Vol1	アクセスID	アクセスID	アクセスID	...
Vol2	アクセスID	アクセスID	アクセスID	...
Vol3	ドライバID	ドライバID	アクセスID	...
Vol4	ドライバID	ドライバID	ドライバID	...

【図 9】

図9

ボリュームID	アクセス要求				...
	リード	ライト	ファイルオープン		
Vol0	許可	許可	許可		...
Vol1	禁止	許可	許可		...
Vol2	禁止	禁止	許可		...
Vol3	——	——	禁止		...

【書類名】 要約書

【要約】

【課題】従来は、高速にホットスタンバイ切替処理を実行する技術を提供することができなかった。

【解決手段】本発明の一実施形態の計算機では、記憶部と、ドライバ手段とアクセス制御手段とを実行する制御部とを有する。ドライバ手段は、システムを起動したときに、共有ディスク装置内の構成情報を取得し、構成情報に基づいて、共有ディスク装置に対するアクセス要求を送信することができる活動状態に設定しておく。また、共有ディスク装置に対するアクセス要求を受信したときに、共有ディスク装置にアクセス要求を送信する。アクセス制御手段は、アクセス先ごとに禁止するアクセス要求を示した管理テーブルに基づいて、アプリケーションプログラムを実行することにより発行されたアクセス要求を送信するかどうかを判定し、禁止されていないアクセス先へのアクセス要求をドライバ手段に送信する。

【選択図】 図 5

認定・付加情報

特許出願の番号	特願 2 0 0 2 - 2 6 6 2 0 6
受付番号	5 0 2 0 1 3 6 3 9 3 3
書類名	特許願
担当官	第七担当上席 0 0 9 6
作成日	平成 1 4 年 9 月 1 3 日

< 認定情報・付加情報 >

【提出日】 平成14年 9月12日

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地

氏 名 株式会社日立製作所